A Novel Hybrid System Based on Fractal Coding for Soccer Retrieval from Video Database

Ehsan Lotfi
Young Researchers Club, Mashhad Branch, Islamic Azad University, Mashhad, Iran.
Email: elotfi@ieee.org

Received: October 2010 Revised: March 2011 Accepted: April 2011

ABSTRACT:
Most information retrieval systems make indirect use of human knowledge in their retrieval process. The novel systems presented here uses the human knowledge directly to retrieve the soccer events. The proposed system is an extended method in three aspects as follows: in the query model, the query by example and query by keywords are applied to retrieve the semantics. In feature extraction, we propose novel methods for extracting the caption as important cinematic feature and for detecting the player gathering. And in the retrieval process, a novel method based on fractal coding is proposed here. The first phase consists of extracting suitable features and key frames from video shots. Then, using a fractal coding, soccer shots are retrieved, and using a fuzzy rule base system, shots that do not include significant soccer events are removed. Experimental results show high accuracy of proposed caption extraction and player gathering detection algorithm and satisfying retrieval process.

KEYWORDS: Soccer, Event Detection, Information Retrieval, Fractal.

1. INTRODUCTION

In many video applications a video conveys a significant message to the audience. For example, in a football match video events such as penalties, corners, goals, and the very concept of victory or loss are conveyed when fans watch the video. The meanings contained in a video are quite obvious for men, but the same does not go for computers. There is a semantic gap between men and machines. For example, machine cannot understand scoring of a goal and corner in a soccer match video, while these meanings are quite obvious for men. Imagine if computers were capable of extracting the meaning of videos then we could utilize them to summarize sport videos automatically [1-3] and retrieving meanings from a multimedia database [4-7]. On the other hand the dramatic increase of multimedia data volumes in recent years will eventually force us to use some sort of retrieval system so that the indexing and retrieval of information handled automatically. The reviewing of retrieval methods from soccer videos may be done in three following scientific areas: the first is the query models. The second, feature extraction and representation methods and the third is indexing methods.

Query model: the user interface of retrieval system is depended on the query model. The queries are done in different manners [8-10]:

1) Query by keywords
2) Sketch-based queries
3) Example-based queries
4) Semantic-based queries

In query by keywords, the user queries the system by using some keywords and the system presents its findings for such ideas as an answer. In sketch-based query or example-based query, the system is queried by user's sketch and example(s). And finally in semantic-based query, the user describes her/his idea and the system must find the idea.

Feature extraction and representation: the extractable features in soccer video can be divided into the two categorizes [11-13]:

1) Object based features and
2) Cinematic features.

For example to recognize the goal we can detect the ball object in the screen and monitor its moving trajectory to be able to recognize the concept of a goal. Such methods use object-based features, and some papers have used this method [1, 5]. On the other hand, certain features may be used to recognize the major events; some of these features are slow motion, whistle, caption, etc. These features are extracted from sound and video sources and are called cinematic features [12-14] divided into two categories:

1) Visual features and
2) Audio features.

Generally if we determine all objects inside the image, including the ball, goals, players, etc. we will be able to recognize many semantics but, this task has high computational complexity. In contrast, using cinematic features provides us with a good tradeoff between the volume of required calculations and preciseness of recognition of high level concepts. In practice, a combination of these methods has been used. And some cases the bit streams of MPEG files have been used [15].

Indexing Methods: Published works include several major methods and techniques for retrieval of information-which are common in some features- as summarized here:

1) Content Based Information Retrieval (CBIR)
2) Semantic Based Retrieval

In CBIR [16-19], media is modeled by features and special relationships between objects and movements. In such methods, suitable features such as color, texture, whistle etc. are extracted and related to high level concepts and meanings, while the system does not know such meanings. In fact the system displays and models the video contents in such a way that is efficient for content retrieval.

For example a statistical model of motion is presented in [10]. After that there is a training session on dynamic contents and then the data are recognized and classified. The result is a database of different video shots (football, basketball, meeting, and highways) each divided into minor events. The provided structure classifies shots in various groups based on similarity of their contents. A fuzzy presentation of video contents described in [20]. In this paper suitable features of a multidimensional fuzzy histogram are made for each video frame. One feature will be suitable for studying the similarities between frames. Consecutive frames are classified based on the fuzzy histogram. In fact the paper divides the videos into shots by this method, and then it chooses a keyframe for each of these frames. Then the shot related to the keyframe is provided as the answer. Zhang et al. [21] proposed a framework for Image retrieval Based on fractal coding which can be used to Concept retrieval [22]. In [21] first, each image in the database is classified by computing information entropy which is compared with a given threshold estimated from the inquired image. Second, the inquired image's fractal codes are generated via Jacquin method. The fractal codes are applied to retrieve images. Finally, the image retrieval result is obtained by matching the similar Euclidean distance between the inquired image and the iterated decoded image.

In other hand, one of the most essential phases of designing a Semantic-based retrieval system is defining a meaningful list of concept-oriented meanings based on human's available knowledge. Thus each single concept in the list should be accompanied by correct and true descriptions on the video collection. A new perspective introduced in recent years includes the use of Semantic Description of Object Motion for the retrieval process, which has been studied in [8], [9], [20], [21] and [23]. The method offered in [8] is as follows: First objects included in video frames are extracted and their trajectories are tracked. Once the trajectories are extracted they are classified for the training of object motion models. Finally, a meaningful description is added to these motion models. It should be added that such a meaningful description was added manually.

2. PROPOSED HYBRID METHOD

The system presented here is for retrieving significant soccer events from database included various videos. The proposed system is improved in three reviewing areas discussed in previous section; the proposed query model is based on both example and text. In the feature extraction, novel methods for caption detection and player gathering detection are proposed. And in indexing step, a novel method based on fuzzy rule based and fractal coding is presented. In the first step the system retrieves soccer shots from database using sample images of soccer provided by user (Fig. 1). In the next step, system retrieves soccer events from shots which are retrieved in the first step. The required input for this step is one or more keywords. The method generally has two steps. In the first step the system retrieves soccer shots from database using fractal coding. The required input for this step is sample images of soccer. In the next step, system extracts the suitable features including captions and player gathering, and retrieves soccer events from shots which are retrieved in the first step. The overall structure is provided in Fig. 1. and Fig. 2 shows the proposed structure in detail. The first phase is devoted to extracting shots from each video and making a list of features extracted from each shot. We used difference of histogram of frames for shot detection. Then a fractal coding stage is used to select soccer video shots and then a list of cinematic features is extracted from the shots. And a fuzzy system is used to eliminate shots including insignificant events. Finally shots are classified and associated with predefined classed using SVM. Then shots related to the class associated with the user query are provided as an answer to that query.

We applied difference of histogram of frames in order to shot boundary detection. After detecting the shots we extract middle frame of each shot as a keyframe. The soccer shots retrieval algorithm is as follow;

Step1. first the input sample image is encoded by Jacquin fractal coding [21-22] with 4*4 fixed child block, and then IFS code of the inquired image's is obtained.

Step2. Current keyframe is as original image for the input sample image IFS code to be decoded with tenth fractal iteration, and fractal decoded image is obtained. The Euclidean distance computed between the decode image and the input sample image is measure of two image similar metric. We choose next keyframe of another shot in the database then goto Step 2.

Step3. These computed Euclidean distances are descending sorted and the smallest or the smaller distance of the first N number keyframes are as the input sample image's same or the similar images. Retrieve shots of N obtained keyframes.

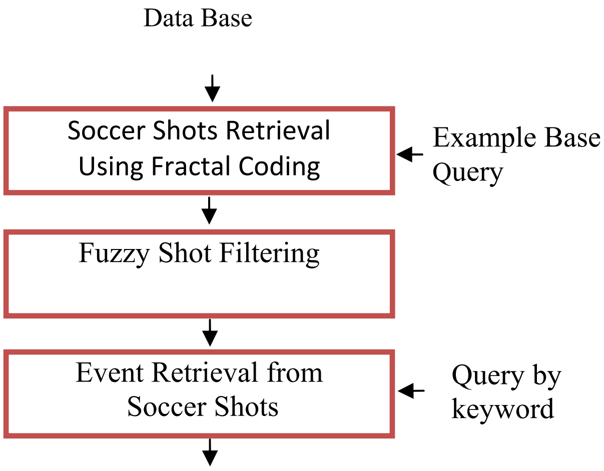


Fig. 1. Retrieval using example based query and query by keyword

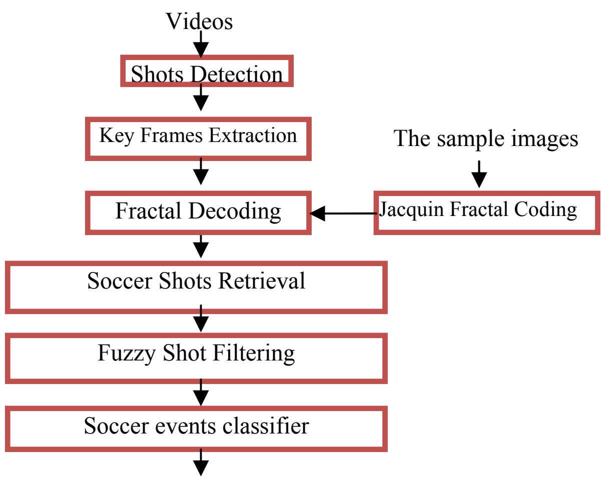


Fig. 2. Feature extraction core and event classification in detail

After retrieving the keyframes of soccer shots the following cinematic features are extracted:

a) Far views from midfield (Fig. 3.a)

b) Far views from field sides (Fig. 3.b)

c) Medium views from inside the field (Fig. 3.c)

d) Outside of the field (Fig. 3.d) and

e) Close-Up (Fig. 3.e).

Also slow-motion replay detection [24] is utilized as a cinematic feature. We have proposed a hierarchical classification method for extracting these views in [25]. The method is presented in Fig. 4 where Gi is percent of grass pixels in keyframe and can be used to classify views (a,b,c) from (d,e). We divided the grass region into 3 sections according to the Figure 5 and calculated G_a and G_b as follows:

$$G_a = \frac{3}{7}G_1 + \frac{4}{7}G_2$$
$$G_b = \frac{3}{7}G_3 + \frac{4}{7}G_2 \qquad (1)$$

where G_i for i =1,2,3 presents the percent of grass pixels in region Ei. G_a and G_b can be used to separate the views (a,b) from (c). By using edge detection algorithm the view (d) from (e) can be classified and by using corner line detection algorithm [26-27], the view (a) from (b) can be categorized.

(a) - (b)
(c) - (d)
(e)

Fig. 3. Types of views: a) Far-center, b) Far-side, c) Medium view, d) Out of field, e) close-up

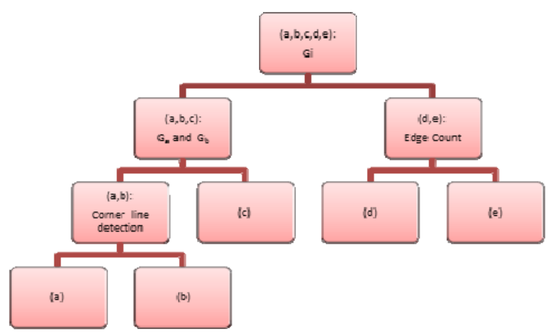


Fig. 4. the hierarchical view classification

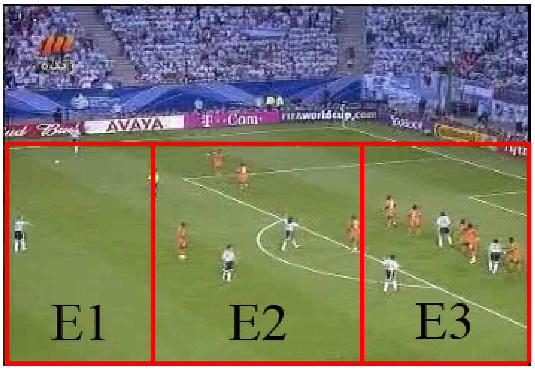


Fig. 5. Segmenting the area including the grass field

In the next section we a novel method for caption extraction and player gathering.

2.1. Novel Feature Extraction Methods

There are two types of caption in a soccer match video: the first is viewed during all or most frames of video. For example the caption which is showing the time and current result of match. The second group is viewed during a few seconds that is depended on events. Here we are focus on second group of captions. Very important features are observed in caption presentation: 1) caption is located at specific area of the picture. For example it's located at the lower half of the picture with specified distances from the sides 2) often it is formed via certain size, direction and shape. For example the best shape to fit the caption is rectangle. 3) The caption is presented during 4 till 7 seconds. 4) The difference between consecutive frames including a caption produces amounts closed to zero at caption location. 5) The caption is quite distinct. By using these features, the following method for caption extraction is proposed here:

Step 1:
-for k = S….S+L, take k th frame
convert to HSI space
for i = 1…m
for j =1..n
$B_k(i,j) = B_k ij,j)+I(k+1)- I(k)$
-for i = 1…m
for j =1..n
if $B_k(i,j) < th1$
$B_k(i,j)=1$
else
$B_k(i,j)=0$
Step 2:
-On the determined search area on matrix B, find the rectangle-shape area including1 and named the caption candidate (C),
-if it is not funded then
f(k) = 0,
S = S+1
Goto the step1
else:
Step 3:
-Convert to the RGB Space

$$-V(R,r) = R^2 + r^2 - \frac{|R+r|^2}{2}$$

$$-V(G,g) = G^2 + g^2 - \frac{|G+g|^2}{2}$$

$$-V(B,b) = B^2 + b^2 - \frac{|B+b|^2}{2}$$

$$-\ f(k) = \frac{V(R,r)+V(G,g)+V(B,b)}{3}$$

-if f(k) > th2 then the founded caption (CC) is valid.
- if S < end of clip then
S = S+1
Goto the Step 1

In the Algorithm, the size of video frames is $m \times n$. *I(k)* is intensity element of *k th* frame. *R,G,B* are the average of colors of red , green and blue of the *k th* frame. And *r,g,b* are the averages of red , green and blue values of the caption candidate area on *k th* frame

Also the search area is determined by observation 2.

1 s on the metrics *B* show the pixels that their gray-levels are not changed during L frames. And *V(k)* is average of variances between caption candidate area and original frame. This measure is defined based on observation 5.

In the far view, a useful feature is player gathering. For example player gathering on the far-side view can show an attack or corner. Here we proposed a simple algorithm to detect player gathering. First by using the grass pixel detection algorithm, the grass area is extracted. By using a threshold on the non-grass pixels count, the player gathering can be detected. Fig. 6 shows the sample result of non-grass detection on specific area.

a)

b)

Fig. 6. The player gathering detection by using player detection on the grass specific area and by using a threshold on the non-grass pixel count

2.2. Using a fuzzy system to eliminate shots containing no significant events

Some of the extracted features as inputs of fuzzy system listed in Table 1 along with abbreviations used for each one of them. Fuzzy system is proposed in order to determine a degree of significance for each shot. For example, the system is expected to give a high degree of significance to a shot which is contained a significant event.

Table 1. The input/output of fuzzy system with their abbreviations

Shot Degree	Out	Sd
Percent of Far-center view in a shot	In	Fc
Percent of Far-side view in a shot	In	Fs
Percent of Medium view	In	Mv
Percent of Out of field	In	Of
Percent of Frame includes Slow-Motion	In	Sm

The system should assign low degree to a shot which is included images of fans and no important event. Rule base and 3 triangle membership functions (low, Medium, High) are presented in Fig. 7 and Fig. 8.

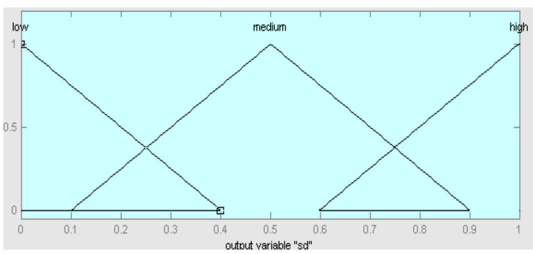


Fig. 7. Triangle membership functions (low, Medium, High) for most In/Out variables

1. If (fc is high) then (sd is low) (1)
2. If (fs is high) then (sd is very high) (1)
3. If (fs is medium) then (sd is high) (1)
4. If (mv is high) then (sd is high) (1)
5. If (of is high) then (sd is low) (1)
6. If (cu is high) then (sd is low) (1)
7. If (cu is medium) then (sd is medium) (1)
8. If (fs is low) then (sd is medium) (1)
9. If (sm is high) then (sd is low) (1)
10. If (sm is medium) then (sd is high) (1)

Fig. 8. Fuzzy rule base for none-useful shot rejection

The inference method is Mamdani product. A sample result of fuzzy inference is shown in Figure 9. So each shot is given a degree of significance by the fuzzy system. The system can reject insignificant shots by thresholding on output of fuzzy system. By defining a threshold(th) on the output of the fuzzy system we can separate the useful and none-useful shots for the purpose of classification. The threshold must be between 0 and 1. The lowest amount of the threshold is zero. By defining the threshold as 0, all of shots in the database will be seen in the output of rejection phase as well. The higher the degree of significance the more sensitive the system vice versa.

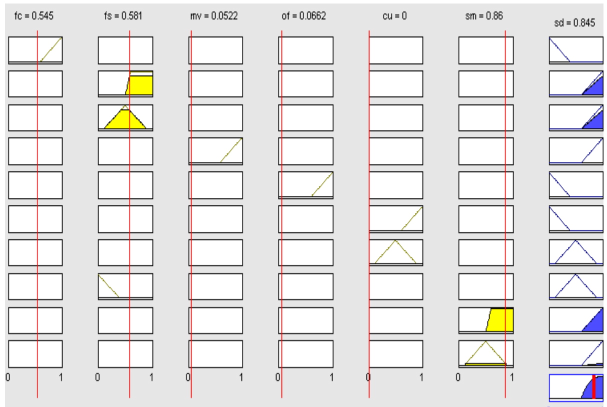


Fig. 9. if (fc=0.545 & fs=0.581 & mv=0.052 & of=0.0682 & cu=0 & sm=0.86) then Shot Degree = 0.845

2.3. Event Retrieval

Up to the previous level the system loaded soccer shots from DB and extracted the features of each shot and then the system eliminated soccer shots including insignificant events. So classification of shots in this phase may be done based on the events included in them. The features listed below are used to classify the events, features include: G_a , G_b, Sd (the output of fuzzy system) and Fs, Percent of Far-side view in a shot. Additionally, Caption (C; 0 or 1) and Player Gathering (P; 0 or 1) are added to feature vector (Fv). After normalizing the above 6 features the value of each of them for each shot will range between 0 and 1, so the feature vector for each shot is shown by Fv (i,V) and expressed in the following form:

$Fv(i,V)=\{G_a , G_b, Fs, C, P, Sd\}$

where: i is the shot number from video V.

We used SVM to classify the events. In the learning phase, events must be associated with 5 predefined classes. All classes have their special keywords which are presented here:

EventClass = { Goal, Penalty, Kroner, Free kick, other}

The provided system will be able to answer questions such as this:

Q1: “Find all the goal shots from all the soccer videos”

3. EXPERIMENTAL RESULTS

We have used more than 130 shots of surveillance videos, shows, sports videos and 119 shots of soccer videos in the database including: 1 match of the World Cup 2006, 2 matches from the UEFA Champions League 2005, 1 match from the FA Premier League 2004, and more than 5 clips from Euro 2004 (Table 2). Sample key frames are shown in Fig. 10. The file format of the input films was non-compressed avi and the size of films is 88×79. Table 3 shows the result of Shot detection.

Table 2. Names and the length (min:sec) of the soccer clips in the database

Euro001(46:40),Euro002(47:00),Euro003(47:30), Euro004(46:00), Euro005(46:45), FIFA-div01f(12:38), FIFA-div02f(18:23),FIFA-div03f(15:38),FIFA-div04f(17:40),FIFA-div05f(18:22),Eng1(42:7) ,Eng2(20:51)

Table 3. The result of Shot detection

Clip Name	Eng2	Euro 005	FIFA-div01f
Length	20:51	18:30	12:35
# of shot	144	133	71
Correct	122	118	65
False	19	11	4
miss	3	4	2
Accuracy(%)	85	89	91

Fig. 10. The key frames of the some video shots in DB

Fig. 11 shows the result of proposed caption extraction algorithm for typical video clip. As illustrated in the figure, *f(k)* resulted from the algorithm is peaked when a caption is appeared. Generally more than 90% captions are detected correctly by using threshold 1000.

And finally in concept retrieval by using tree sequences illustrated in Table 1 more than 90% goals and Corners are extracted.

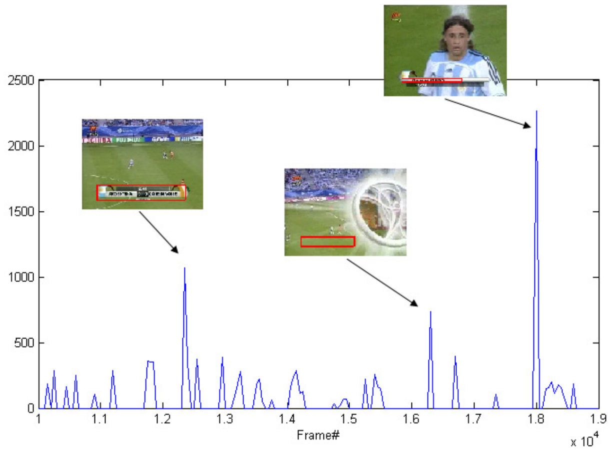


Fig. 11. Caption extraction results on typical video clip. the *f(frame#)* can extract the captions.

Table 4. The result of example based query in soccer shot retrieval step using fractal coding

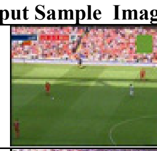
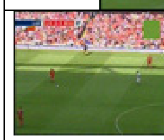
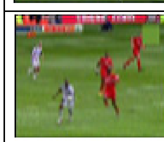
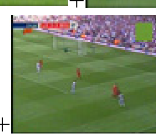
Input Sample Image(s)	**Correct**	**False**	**Accuracy**
	35	12	29%
	75	26	63%
 + 	93	31	78%
 +  + 	102	33	86%

Table 4 shows the result of soccer shot retrieval using fractal coding. Each row presents the result of one or more sample input image. In this table accuracy was calculated by

Accuracy = 100* Correct/119.

The table indicate the higher the number of sample images, the higher the accuracy of retrieval process.

In the training phase we used 119 shots to learning for SVM and the rate of correct class in training data is 87%. As noted earlier the output of fuzzy system is a degree for each shot and the system rejects no useful shots using a threshold (th) as system input parameter. The input margin of the system for evaluation is 0.2 and then inputs include 0.2, 0.4, 0.6, 0.8 and 1. The results of changing the input amount of the system and the accuracy of the events classification are shown in tables 5,6,7,8. In this evaluation scenario we used 187 shots include 5 Goals, 5 Penalties, 6 Free kicks, 9 Corners and 162 shots containing no significant events. Table 5 shows that the th is 0.2 then system reject 23 shots. In table 6 where the th is 0.4 then system reject 86 shots. In table 7 where th is 0.6 then the system reject 114 shots and in table 8 where th is 0.8 then the system reject 172 shots.

Table 5. The result of proposed method with th = 0.2

	Goal	Penalties	Free kick	Corner	other	sum
Total	5	5	6	9	162	187
# of rejected	0	0	0	0	23	23
Correct	5	5	6	9	86	111
False	13	9	17	14	0	53
Accuracy	27%	36%	26%	39%	-	-

Table 6. The result of proposed method with th = 0.4

	Goal	Penalties	Free kicks	Corner	other	sum
Total	5	5	6	9	162	187
# of rejected	0	0	0	0	86	86
Correct	5	5	5	8	51	74
False	6	0	13	6	1	27
Accuracy	45%	100%	32%	53%	-	-

Table 7. The result of proposed method with th = 0.6

	Goal	Penalties	Free kicks	Corner	other	sum
Total	5	5	6	9	162	187
# of rejected	0	0	2	2	110	114
Correct	5	5	4	7	42	63
False	2	0	5	3	0	10
Accuracy	71%	100%	36%	58%	-	-

In table 7 where th is 0.6, 4 shots containing corner and free kicks are rejected incorrectly (see the '# of rejected' row) and in table 8 where th is 0.8, 10 shots containing very important events are rejected incorrectly.

In these tables accuracy is calculated by:

Accuracy =100* Correct/(Total + False)

The tables indicate that the higher the threshold, the higher the number of rejected shots will be. Also, the higher the threshold, the lower the number of incorrect shot detection will be (see 'False' row of tables). See Table 8 where th = 0.8 and False detection for all events is zero but 1 Goal, 5 Free kicks and 4 corner are rejected incorrectly.

Table 8. The result of proposed method with th = 0.8

	Goal	Penalties	Free kicks	Corner	other	sum
Total	5	5	6	9	162	187
# of rejected	1	0	5	4	162	172
Correct	4	5	1	5	0	15
False	0	0	0	0	0	0
Accuracy	80%	100%	16%	55%	-	-

4. CONCLUSIONS

The proposed method uses fractal coding for soccer shots retrieval from video databases and uses a fuzzy rule base containing the experiences of experts, to reject insignificant soccer shots. Additionally, in this paper, novel methods for caption detection and player gathering detection are proposed. According to the experimental results, the proposed feature extraction methods can extract more than 90% of captions in soccer video clips. And proposed soccer retrieval system can extract the significant events.

5. ACKNOWLEDGMENT

This work is supported by Young Researchers Club, Mashhad Branch of Islamic Azad University. Special thanks to Dr. Mohsen Allame, the chairman of the club and to Mr. Hesami for his feedback on the final version of the paper.

REFERENCES

[1] Maheshkumar H. Kolekar, Kannappan Palaniappan, **"Semantic concept mining based on hierarchical event detection for soccer video indexing,"** *Journal of multimedia*, Vol. 4, No. 5, October 2009.

[2] Ahmet Ekin, A. Murat Tekalp and Rajiv Mehrotra, **"Automatic soccer video analysis and summarization,"** *IEEE Transactions on image processing*, Vol. 12, No. 7, July 2003.

[3] Kiani, V., Pourreza, H.R., **"Robust GME in encoded mpeg video,"** *Proceeding of the 9th International Conference on Advances in Mobile Computing and Multimedia (MoMM 2011)*, ACM, pp. 147-154.

[4] Sun Xing-hua, Yang Jing-yu, **"Inference and retrieval of soccer event,"** *Mar. 2007, Volume 4, No.3 (Serial No.28) Journal of Communication and Computer*.

[5] Changsheng Xu ,Jinjun Wang, Hanqing Lu, and Yifan Zhang, **"A novel framework for semantic annotation and personalized retrieval of sports video"**, *IEEE transactions on multimedia,* Vol. 10, No. 3, April 2008.

[6] T. D'Orazio , M. Leo, P. Spagnolo, M. Nitti, N. Mosca, A. Distante, **"A visual system for real time detection of goal events during soccer matches,"** *Computer Vision and Image Understanding 113* (2009) 622–632.

[7] T. D'Orazio, M. Leo, P. Spagnolo, P. L. Mazzeo, N. Mosca, M. Nitti, A. Distante, **"An investigation into the feasibility of real-time soccer offside detection from a multiple camera system"** *in Proc*, 2009.

[8] Yu-Lin Kang, **"Soccer video event detection with visual keywords"**, *ICICS-CM 200315-18 December 2003*, Singapore.

[9] Z. Q. Liu, L. T. Bruton, J. C. Bezdek, J. M. Keller, S. Dance, N. R. Bartley, and C. Zhang, **"Dynamic image sequence analysis using fuzzy measures,"** *IEEE Trans. Syst., Man, Cybern.*, Vol. 31, No. 4, pp. 557–571, Aug. 2001.

[10] Weiming Hu, Dan Xie, Zhouyu Fu, Wenrong Zeng, and Steve Maybank, **"Semantic-based surveillance video retrieval"**, *IEEE Transactions on image processing*, Vol. 16, No. 4, April 2007.

[11] Liang Bai, Songyang Lao, Weiming Zhang, **"A semantic event detection approach for soccer video based on perception concepts and finite state machines,"** *Eight International Workshop on Image Analysis for Multimedia Interactive Services, WIAMIS'07.*

[12] Ahmet Ekin, A. Murat Tekalp and Rajiv Mehrotra, **"Automatic soccer video analysis and summarization"**, *IEEE Transactions on image processing,* Vol. 12, No. 7, July 2003.

[13] T.D'Orazio, M. Leo, **"A review of vision-based systems for soccer vide analysis,"** *PatternRecognition, doi:10.1016/j.patcog.* February2, 2010

[14] Shu-Ching Chen, Mei-Ling Shyu, Chengcui Zhang, Lin Luo, Min Chen **"Detection of soccer goal shots using joint multimedia features and classification rules,"** *MDM/KDD'03*, August 27, 2003, Washington, DC, USA.

[15] R. Leonardi, P. Migliorati and M. Parandini, **"Semantic indexing of sports program sequences by audio-visual analysis,"** *in ICIP Conf.*, 2003.

[16] Geetha, Vasumathi Narayanan, **"A survey of content-based video retrieval,"** *Journal of Computer Science 4 (6):* 474-486, 2008, 1549-3636.

[17] David Liu , Tsuhan Chen, **"Video retrieval based on object discovery,"** *Computer Vision and Image Understanding 113* (2009) 397–404.

[18] Jinhui Tang a, Guo-JunQi, MengWang , Xian-ShengHua**, "Video semantic analysis based on structure-sensitive anisotropic manifold ranking ,"** *Signal Processing 89* (2009) 2313–2323.

[19] Rong Yan Æ Alexander G. Hauptmann, **"A review of text and image retrieval approaches for broadcast news video,"** *Inf Retrieval* (2007) 10:445–484.

[20] A. D. Doulamis, N. D. Doulamis and S. D. Kollias, **"A fuzzy video content representation for video summarization and content-based retrieval,"** *Signal Process.* 80 (2000) 1049–1067.

[21] L. Zhang, L. Xi and B. Zhou, **"Image retrieval method based on entropy and fractal coding,"** *WSEAS Trans. Syst*. 4(7) (2008) 232–241.

[22] E. Lotfi, Yaghoobi, M., **"Concept retrieval based on a combination of fractal coding, fuzzy rule based system and SVM, "** *FRACTALS , Complex Geometry, Patterns, and Scaling in Nature and Society*, Volume: 19, Issue: 2(2011) pp. 185-194 DOI: 10.1142/S0218348X11005300

[23] G. Piriou, P. Bouthemy and J.-F. Yao, **"Recognition of dynamic video contents with global probabilistic models of visual motion,"** *IEEE Trans. Image Process*. 15(11) (2006) 3418–3431.

[24] H. Pan, P. van Beek, and M. I. Sezan, **"Detection of slow-motion replay segments in sports video for highlights generation,"** *in Proc. IEEE Int'l. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, 2001.

[25] E. Lotfi and H. R. Poureza, **"Summarizing soccer videos without detecting the events,"** *Inernational Journal of Computer and Netork Security* 6(10) (2006) 130–135.

[26] Ehsan Lotfi, Hamid Reza Pourreza, **"Event detection and automatic summarization in soccer video,"** *MVIP2007*.

[27] Ehsan Lotfi, M.R. Akbarzadeh-T. and H.R. Pourreza, **"A hybrid method for soccer video events retrieval using fuzzy systems,"** *International Fuzzy Systems Association Conference (IFSA2009)*, Lisbon, July 2009.